SILAN – AI REAL-TIME SIGN LANGUAGE DETECTION

Submitted for
CSET301: Artificial Intelligence and Machine Learning

Submitted by:

V S SANTHOSH (E23CSEU1150)
BHANUTEJA PADAMATA (E23CSEU1144)
SHUBHAM PANDEY (E23CSEU1167)

Submitted to
SHWETANG DUBEY

JAN-APRIL 2025
SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

INDEX

Sr. No	Content	Page No
1.	Abstract	1
2.	Introduction	2
3.	Related Work	3
4.	Problem Statement	4
5.	Methodology	6
6.	Contribution	7
7.	Result	8
8.	Conclusions	9
9.	References	10

Abstract

SILAN (Sign Language Interpreter using AI in real-time) is a powerful deep learning project aimed at bridging the communication gap between the hearing and speech-impaired community and the rest of the world. The system leverages Convolutional Neural Networks (CNNs) and real-time computer vision using OpenCV to detect and classify hand gestures corresponding to letters in Indian Sign Language. Implemented using both TensorFlow and PyTorch for comparative learning and flexibility, this project focuses on capturing live video frames, processing hand gesture inputs, and outputting translated text in real-time.

To enhance usability, SILAN is wrapped in a Flask-based full-stack web application, allowing users to access sign language translation from their browser. Extensive training was carried out on custom datasets with data augmentation to improve accuracy and generalization. This report provides a thorough analysis of model architecture, performance benchmarks, and real-time deployment. SILAN exemplifies how AI can empower inclusive technology and contribute meaningfully to accessibility.

Introduction

Communication is a fundamental human need, yet millions of individuals with hearing or speech impairments face challenges in everyday interactions. Sign Language serves as a bridge, but its usage is often limited by the lack of widespread understanding among the general population. Recent advancements in AI and computer vision provide an opportunity to build systems that can understand and interpret sign language, facilitating seamless communication.

This project introduces SILAN, a real-time sign language detection system built using deep learning and computer vision. It employs pre-trained CNN architectures and hand gesture recognition to classify alphabet signs from Indian Sign Language (ISL). The integration of real-time camera input through OpenCV and a simple web interface via Flask ensures that the solution is practical, accessible, and responsive.

Through comparative implementation in TensorFlow and PyTorch, learners can grasp the core concepts and architectural choices in deep learning. By turning theory into a usable tool, SILAN embodies the potential of AI in social good and inclusive design.

Related Work

Various approaches to Sign Language Recognition (SLR) have evolved over time:

- DeepHand and SignNet: CNN-based architectures for static sign detection.
- Mediapipe by Google: Provides hand tracking using landmarks, often used for gesture-based recognition systems.
- ASL Recognition using CNN + LSTM: Combines spatial and temporal features for dynamic sign language recognition.
- Real-Time Sign Recognition with OpenCV and TensorFlow: Simple but effective pipeline for capturing gestures via webcam and predicting classes.

These frameworks and techniques laid the foundation for our custom approach in SILAN, emphasizing real-time detection and deployment.

Problem Statement

To develop a real-time sign language interpreter that:
- Accurately classifies hand gestures from video feed using CNN.
- Supports ISL alphabets through image recognition.
- Integrates both TensorFlow and PyTorch implementations for learning and benchmarking.
- Deploys a user-friendly web application using Flask for seamless interaction.
- Addresses challenges like background noise, gesture variance, and lighting conditions.

Methodology

1. Data Collection & Preprocessing
- Custom dataset of ISL hand signs (A-Z).
- Data augmentation (rotation, scaling, flipping).
- Normalization and resizing to 64x64.

2. Model Architecture
- CNN with multiple convolutional and pooling layers.
- Dropout for regularization.
- Softmax classifier for output.

3. PyTorch Pipeline
- Dataset and Dataloader setup.
- Model training using cross-entropy loss and Adam optimizer.
- Save model as .pt file for deployment.

4. TensorFlow Pipeline
- Sequential API with Conv2D and MaxPooling layers.
- Training loop with validation metrics.
- Export as .h5 file.

5. Web Deployment with Flask
- Webcam input via OpenCV.
- Live prediction overlay using model inference.
- UI for uploading or live detection with results shown on screen.

Contribution

- Dual Deep Learning Implementation: TensorFlow and PyTorch versions.
- Flask Web App: Real-time gesture detection through browser interface.
- Real-World Dataset Creation: Custom ISL dataset to improve accuracy.
- Educational Resource: Helps learners understand CNNs, deployment, and computer vision.

Result

- Achieved 63%+ accuracy on test data.
- Real-time detection with minimal latency.
- Robust across varying lighting and hand orientations.
- Web interface runs on local server and supports video gesture input.
- Models compared across frameworks for training time, accuracy, and resource usage.

Conclusions

SILAN showcases the potential of artificial intelligence in building inclusive and impactful solutions. By recognizing Indian Sign Language gestures in real time, it bridges a crucial gap in communication. With dual framework implementation, an interactive web app, and high accuracy, this project is both a technical achievement and a social innovation. Future improvements include dynamic gesture support, multi-hand recognition, and multilingual translation.

References

1. OpenCV Documentation – https://docs.opencv.org
2. TensorFlow Guide – https://www.tensorflow.org
3. PyTorch Docs – https://pytorch.org/docs/
4. Mediapipe by Google – https://google.github.io/mediapipe/
5. Indian Sign Language Dataset Resources
6. Flask Documentation – https://flask.palletsprojects.com/
7. CNN for Hand Gesture Recognition – Research papers and blog implementations